

STIC-Biotech/ChemLib

76379

From: Fredman, Jeffrey
Sent: Tuesday, September 24, 2002 10:20 AM
To: STIC-Biotech/ChemLib
Cc: Schmidt, Mary
Subject: FW: sequence search request 09/923,831

I Approve.

Please Rush.

Jeff Fredman

-----Original Message-----

From: Schmidt, Mary
Sent: Tuesday, September 24, 2002 9:53 AM
To: Fredman, Jeffrey
Subject: RE: sequence search request 09/923,831

2365 bases

-----Original Message-----

Fr m: Fredman, Jeffrey
Sent: Tuesday, September 24, 2002 7:06 AM
To: Schmidt, Mary
Subject: RE: sequence search request 09/923,831

How big is the sequence?

-----Original Message-----

From: Schmidt, Mary
Sent: Monday, September 23, 2002 3:33 PM
To: Fredman, Jeffrey
Subject: FW: sequence search request 09/923,831

Hi, I ordered a full length search of SEQ ID NO:42 last month and got the results, however, I didn't notice then that they also have fragment claims-- could you please rush the following request?

Please search SEQ ID NO:42: (1) results size limited to less than 100 bases, and (2) results size limited from 100-
~~300 bases~~. (NO interference search and NO full-length search).

Thanks!!

Melissa
11D05
mailbox 11E12
au. 1635

RECEIVED
SEP 24 2002
STIC Division



We can only
limit and
not range

So, there is
100 & 300
only

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 9/24/02
Date Completed: _____
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, **.rnpm** and **.rnpn**

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, **.rapm** and **.rapn**

The Pending database search results should not be left in the case because they contain data that is confidential.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 19:15:48 ; Search time 4266.96 Seconds
(without alignments)
11598.716 Million cell updates/sec

Title: US-09-923-831-42
Perfect score: 2365
Sequence: 1 ttgtaccgagctcgatcc.....aaaaaaaaaaaaaaaaaa 2365

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 1365494

Minimum DB seq length: 0
Maximum DB seq length: 300

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
------------	-----------	-------	-------	--------	-------	-------------

1	62	2.6	234	6	AX314960
2	59.8	2.5	229	11	G66079
3	50.8	2.1	235	1	STREXP98
4	50.8	2.1	235	6	AR084893
5	47	2.0	224	3	DDIACD
6	47	2.0	276	6	AX308194
7	46.8	2.0	247	6	AX188933
8	46.8	2.0	280	8	AB027882
9	46.4	2.0	268	6	AX185659
10	44.4	1.9	220	6	AX185507
11	44.4	1.9	300	6	AX185628
12	44	1.9	160	3	DDIACD
13	43.6	1.8	294	6	AX185661
14	43.4	1.8	140	3	DMRRL
15	43	1.8	288	6	AX185400
16	42.4	1.8	220	11	G58490
17	42.2	1.8	174	6	AX340337
18	41.8	1.8	289	6	AX198818
19	41.6	1.8	229	6	AX062568
20	41.4	1.8	255	6	AX210942
21	41.4	1.8	272	3	ECA270206
22	41.2	1.7	255	6	AX211073
23	41	1.7	116	6	AX08900
24	41	1.7	119	6	AX08901
25	41	1.7	119	6	AX08904
26	41	1.7	198	6	AX182091
27	41	1.7	202	6	AX099441
28	41	1.7	206	10	AF168586
29	40.8	1.7	250	6	AX210706
30	40.6	1.7	192	11	G38017
31	40.6	1.7	263	11	CNS06EHA
32	40.4	1.7	226	6	AX184991
33	40.2	1.7	190	11	HUMT5193A
34	40.2	1.7	252	13	AF303484
35	40	1.7	143	6	AX255922
36	39.8	1.7	233	6	AX182078
37	39.8	1.7	240	8	MITGIR1
38	39.6	1.7	112	6	AX261671
39	39.6	1.7	202	4	DOGSNVD17B
40	39.4	1.7	99	6	AX08898
41	39.4	1.7	104	6	AX08899
42	39.4	1.7	222	6	AX122013
43	39.4	1.7	252	6	AX197621
44	39.4	1.7	258	10	BC014856
45	39.4	1.7	295	6	AX184943

ALIGNMENTS

RESULT 1	AX314960	Sequence 7945 from Patent WO0190366.	234 bp	DNA	linear	PAT 14-DEC-2001
LOCUS	AX314960					
DEFINITION	AX314960					
ACCESSION	AX314960					
VERSION	AX314960.1	GI:17898715				
KEYWORDS						
SOURCE	human.					
ORGANISM	Homo sapiens					

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Leach,M.D. and Shinkets,R.A.
Human polynucleotides and polypeptides encoded thereby
Patent: WO 0190366-A 7945 29-NOV-2001;
Curagen Corporation (US)
Location/Qualifiers
1. .234
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT	59 a	60 c	54 g	61 t
ORIGIN				


```

/db_xref="GI:398103"
/translation="AFVFGTRRRRVDELFRGLKIRGFRAESIGDLDQNKRLRVLRD"
KNLNDLVADVAARGLDISGTVHVNYPD"
BASE COUNT      58 a      48 c      58 g      71 t
ORIGIN

Query Match          2.1%; Score 50.8; DB 1; Length 235;
Best Local Similarity 51.3%; Pred. No. 1.3;
Matches 118; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 1679 ttgtcttcgtttctcgaaagcgtgttcggatcaacttataagtgacctaaacttgaa 1738
Db      5 TTGATTTTGGTCGACCAACCGCGTGTGATGAATGACTCGTGGTTTGAATAATTCGT 64
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1739 atatatcagtagctctctgcattgagatagaaacagagatcgggagaaagcattag 1798
Db      65 GCTTCGTCGACAGAAGGAATTCATGCGACCTAGACCAAAACGCTCTTCGTGCTTC 124
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1799 agaactttaaacacaggcaagtgaatactaattgcaactgatctagcctctagagac 1858
Db      125 GTGACTTTAAATATGGCANFTCTGATGTTTGTGTGGACAGAGCTGCAGCGCTGTT 184
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1859 ttgatgtccatgaagttaacatgctctataattttgaactttccacgaat 1908
Db      185 TGGATATTCAGGTGTGACCAATGCTCAACACTACGATATTCACAGAT 234
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
LOCUS      AR084893              235 bp      DNA      linear      PAT 01-SEP-2000
DEFINITION Sequence 19 from patent US 5981229.
ACCESSION  AR084893
VERSION    AR084893.1 GI:10011664
KEYWORDS
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 235)
AUTHORS   Measure,H.Robert, Pearce,B.J. and Tuomanen,E.
TITLE     Bacterial exported proteins and acellular vaccines based thereon
JOURNAL   Patent: US 5981229-A 19 09-NOV-1999;
FEATURES   Location/Qualifiers
            source          1..235
                        /organism="unknown"

BASE COUNT      58 a      48 c      58 g      71 t
ORIGIN

Query Match          2.1%; Score 50.8; DB 6; Length 235;
Best Local Similarity 51.3%; Pred. No. 1.3;
Matches 118; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 1679 ttgtcttcgtttctcgaaagcgtgttcggatcaacttataagtgacctaaacttgaa 1738
Db      5 TTGATTTTGGTCGACCAACCGCGTGTGATGAATGACTCGTGGTTTGAATAATTCGT 64
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1739 atatatcagtagctctctgcattgagatagaaacagagatcgggagaaagcattag 1798
Db      65 GCTTCGTCGACAGAAGGAATTCATGCGACCTAGACCAAAACGCTCTTCGTGCTTC 124
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1799 agaactttaaacacaggcaagtgaatactaattgcaactgatctagcctctagagac 1858
Db      125 GTGACTTTAAATATGGCANFTCTGATGTTTGTGTGGACAGAGCTGCAGCGCTGTT 184
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1859 ttgatgtccatgaagttaacatgctctataattttgaactttccacgaat 1908
Db      185 TGGATATTCAGGTGTGACCAATGCTCAACACTACGATATTCACAGAT 234
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 5
LOCUS      DDIACTD              224 bp      DNA      linear      INV 27-APR-1993
DEFINITION

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.
OM nucleic - nucleic search, using sw model
Run on: September 24, 2002, 19:56:58 ; Search time 371.69 Seconds
(without alignments)
10924.433 Million cell updates/sec

Title: US-09-923-831-42
Perfect score: 2365
Sequence: 1 ttgtaccgagctcgatcc.....aaaaaaaaaaaaaaaa 2365
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 2377130

Minimum DB seq length: 0
Maximum DB seq length: 300

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	139	5.9 139 22	AAK24181 Human brain expres
2	130	5.5 130 22	ABA74723 Human foetal liver
3	130	5.5 130 22	ABA39434 Probe #17900 for g
4	130	5.5 130 22	AAK23210 Human brain expres
5	130	5.5 130 22	AAK43982 Human bone marrow
6	130	5.5 130 22	AAI26493 Probe #16426 for g
7	130	5.5 130 22	AAI55243 Probe #23929 used
8	53.4	2.3 259 21	AAA42361 Human secreted exp
9	53.4	2.3 299 20	AAK84188 DNA encoding human

10	53.4	2.3	299	21	AACT9417	5' cDNA sequence o
11	50.8	2.1	235	16	AAQ83249	Streptococcus pneu
12	47.2	2.0	204	15	AAQ57421	DEAD ATP helicase
13	47	2.0	102	24	ABK09532	Human ovarian tumo
14	46.8	2.0	197	23	AAK49363	Staphylococcus aur
15	46.8	2.0	247	22	AAH71353	Human cervical can
16	46.4	2.0	268	22	AAH70080	Human cervical can
17	45.2	1.9	282	21	AAH08972	Fusarium venenatum
18	44.4	1.9	220	22	AAH69928	Human cervical can
19	44.4	1.9	300	22	AAH70043	Human cervical can
20	44.2	1.9	165	22	AAI84674	Human polynucleoti
21	44	1.9	232	22	AAH33958	Human colon cancer
22	43.8	1.9	293	22	AAI25205	Human breast cancer
23	43.6	1.8	213	15	AAQ57415	Human breast cancer
24	43.6	1.8	289	22	AAI16073	Human breast cancer
25	43.4	1.8	234	22	AAH70082	Human cervical can
26	43.2	1.8	259	18	AAH29038	cDNA encoding for
27	43.2	1.8	259	18	AAH29038	Human H1075-1 secr
28	43	1.8	225	19	AAV26699	Human novel secret
29	43	1.8	225	21	AAA40538	Human adult ovary
30	43	1.8	288	22	AAH69821	Human cervical can
31	42.6	1.8	191	17	AAH13475	Capture probe for
32	42.6	1.8	259	18	AAV00423	3' fragment of clo
33	42.4	1.8	236	22	AAI16239	Human breast cancer
34	42.2	1.8	271	22	AAI16362	Human breast cancer
35	42.2	1.8	296	22	AAI09995	Human breast cancer
36	42.2	1.8	296	22	AAI19575	Human breast cancer
37	42	1.8	285	22	AAI34854	Human musculoskele
38	41.8	1.8	178	22	AAH24361	Retroviral recombi
39	41.8	1.8	208	22	AAH24363	Retroviral recombi
40	41.8	1.8	289	22	AAH63649	Human ovarian tumo
41	41.8	1.8	295	22	AAH29032	cDNA encoding for
42	41.6	1.8	233	22	AAH68277	Human lung tumour
43	41.4	1.8	230	22	AAI00088	Human reproductive
44	41.4	1.8	255	22	AAH82075	Rat differential t
45	41.4	1.8	299	21	AAA00517	Human colon cancer

ALIGNMENTS

RESULT 1
AAK24181/c
ID AAK24181 standard; DNA; 139 BP.
XX
AC AAK24181;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 24172.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-483446/52.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 PS Example 4; SEQ ID NO: 24172; 650pp + Sequence Listing; English.
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.
 XX Sequence 139 BP; 34 A; 33 C; 27 G; 45 T; 0 other;
 SQ

Query Match 5.9%; Score 139; DB 22; Length 139;
 Best Local Similarity 100.0%; Pred. No. 3.7e-21;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1814 gcaagtgaatactaatgcaactgattctagctctagagacttgatctccatgacy 1873
 bb 139 GCAAGTGAATACTAATTGCACTGATCTAGCCTCTAGAGACTTGATGTCATGACG 80
 QY 1874 ttacacatgtctataatttgactttccacggaatattggaataatcgtacacgaatag 1933
 db 79 TTACATGCTCTATAATTGACTTTCCACGGAATATTGAGAAATACGTACACCGAATAG 20
 QY 1934 ggcgcacgggaagacgagc 1952
 db 19 GGCACGCGGAGACGAGG 1

RESULT 2
 ABA74723
 ID ABA74723 standard; DNA; 130 BP.
 XX
 AC ABA74723;
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Human foetal liver single exon nucleic acid probe #23028.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 XX
 PR 26-MAY-2000; 2000US-0207456.
 XX
 PR 30-JUN-2000; 2000US-0608408.
 XX
 PR 03-AUG-2000; 2000US-0632366.
 XX
 PR 21-SEP-2000; 2000US-0234687.
 XX
 PR 27-SEP-2000; 2000US-0236359.
 XX
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 XX WPI; 2001-483447/52.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 PS

XX Claim 4; SEQ ID NO 23028; 639pp + sequence listing; English.
 PS
 XX The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 130 BP; 60 A; 18 C; 24 G; 28 T; 0 other;
 XX

Query Match 5.5%; Score 130; DB 22; Length 130;
 Best Local Similarity 100.0%; Pred. No. 3.4e-19;
 Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 514 ataatacaagaacaacacagaatcattagtcacaaatttttggcagcaagcaatgcaaacg 573
 db 1 ataatacaagaacaacacagaatcattagtcacaaatttttggcagcaagcaatgcaaacg 60
 QY 574 aaagcaaaagcagtgatagacaattttgttaaagctagaagaaaattacaattcagaa 633
 db 61 aaagcaaaagcagtgatagacaattttgttaaagctagaagaaaattacaattcagaa 120
 QY 634 tgcgggaattg 643
 db 121 tgcgggaattg 130

RESULT 3
 ABA39434
 ID ABA39434 standard; DNA; 130 BP.
 XX
 AC ABA39434;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Probe #17900 for gene expression analysis in human heart cell sample.
 XX
 KW Human; gene expression; heart; microarray; vascular system; probe;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157274-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00666.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 XX
 PR 26-MAY-2000; 2000US-0207456.
 XX
 PR 30-JUN-2000; 2000US-0608408.
 XX
 PR 03-AUG-2000; 2000US-0632366.
 XX
 PR 21-SEP-2000; 2000US-0234687.
 XX
 PR 27-SEP-2000; 2000US-0236359.
 XX
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 XX WPI; 2001-488899/53.
 XX
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 PS Claim 4; SEQ ID No 17900; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart. The
 CC present sequence is one such probe. The probes may be used for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from the human heart via microarrays. By measuring gene expression, the
 CC probes are useful for predicting, diagnosing, grading, staging,
 CC monitoring and prognosing diseases of the human heart and vascular system
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
 CC congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 130 BP; 60 A; 18 C; 24 G; 28 T; 0 other;

Query Match 5.5%; Score 130; DB 22; Length 130;
 Best Local Similarity 100.0%; Pred. No. 3.4e-19;
 Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 514 ataatacaagaacacaccagaatcattagtcaaaatttttggcagcgaagcgaatgcaaacg 573
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 ataatacaagaacacaccagaatcattagtcaaaatttttggcagcgaagcgaatgcaaacg 60
 QY 574 aaagcaaaagcagtgatagacaattttgttaaaagctagaagaataattacaaattcgaa 633
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 aaagcaaaagcagtgatagacaattttgttaaaagctagaagaataattacaaattcgaa 120
 QY 634 tgcggaattg 643
 ||||||||
 Db 121 tgcggaattg 130

RESULT 4
 AAK23210
 ID AAK23210 standard; DNA; 130 BP.
 XX
 XX AAK23210;
 XX
 DT 05-NOV-2001, (first entry)
 XX
 DE Human brain expressed single exon probe SEQ ID NO: 23201.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer; ss.
 XX
 OS Homo sapiens.
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DX, Chen W, Rank DR;
 XX
 DR WPT; 2001-483446/52.
 XX

PT Single exon nucleic acid probes for analyzing gene expression in human
 XX brains -
 XX

PS Example 4; SEQ ID NO: 23201; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.
 XX
 SQ Sequence 130 BP; 60 A; 18 C; 24 G; 28 T; 0 other;

Query Match 5.5%; Score 130; DB 22; Length 130;
 Best Local Similarity 100.0%; Pred. No. 3.4e-19;
 Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 514 ataatacaagaacacaccagaatcattagtcaaaatttttggcagcgaagcgaatgcaaacg 573
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 ataatacaagaacacaccagaatcattagtcaaaatttttggcagcgaagcgaatgcaaacg 60
 QY 574 aaagcaaaagcagtgatagacaattttgttaaaagctagaagaataattacaaattcgaa 633
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 aaagcaaaagcagtgatagacaattttgttaaaagctagaagaataattacaaattcgaa 120
 QY 634 tgcggaattg 643
 ||||||||
 Db 121 tgcggaattg 130

RESULT 5
 AAK49382
 ID AAK49382 standard; DNA; 130 BP.
 XX
 XX AAK49382;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed single exon probe SEQ ID NO: 23939.
 XX
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
 XX
 OS Homo sapiens.
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00668.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DX, Chen W, Rank DR;
 XX
 DR WPT; 2001-488900/53.
 XX

PT Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human bone marrow -

PS Example 4; SEQ ID NO: 23939; 658pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is one of
 CC the probes of the invention.

XX
 SQ Sequence 130 BP; 60 A; 18 C; 24 G; 28 T; 0 other;

Query Match 5.5%; Score 130; DB 22; Length 130;
 Best Local Similarity 100.0%; Pred. No. 3.4e-19;
 Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 514 ataatacagaacaccagaaatcattagtcacaaatttttggcagcaggaatgcaaacg 573
 |||||
 Db 1 ataatacagaacaccagaaatcattagtcacaaatttttggcagcaggaatgcaaacg 60
 QY 574 aaagcaaaagcagtgatagacaaattttgttaaaagctagagaagaaattacaaattcgaa 633
 |||||
 Db 61 aaagcaaaagcagtgatagacaaattttgttaaaagctagagaagaaattacaaattcgaa 120
 QY 634 tgcggaattg 643
 |||||
 Db 121 tgcggaattg 130

RESULT 6

AAI26493
 ID AAI26493 standard; DNA; 130 BP.

XX AC AAI26493;
 XX DT 12-OCT-2001 (first entry)
 DE Probe #16426 for gene expression analysis in human cervical cell sample.
 DE Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer; ss.
 XX OS Homo sapiens.

XX PN WO200157278-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00670.
 XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488901/53.
 XX DR Human genome-derived single exon nucleic acid probes useful for
 XX PT analyzing gene expression in human cervical epithelial cells -
 XX PS Claim 25; SEQ ID No 16426; 487bp; English.

XX CC The present invention relates to human single exon nucleic acid probes
 CC (SENP). The present sequence is one such probe. The SENPs are derived
 CC from human hela cells. The SENPs can be used to produce a single exon
 CC microarray, which can be used for measuring human gene expression in a
 CC sample derived from human cervical epithelial cells. By measuring gene
 CC expression, the probes are therefore useful in grading and/or staging
 CC of diseases of the cervix, notably cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 130 BP; 60 A; 18 C; 24 G; 28 T; 0 other;

Query Match 5.5%; Score 130; DB 22; Length 130;
 Best Local Similarity 100.0%; Pred. No. 3.4e-19;
 Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 514 ataatacagaacaccagaaatcattagtcacaaatttttggcagcaggaatgcaaacg 573
 |||||
 Db 1 ataatacagaacaccagaaatcattagtcacaaatttttggcagcaggaatgcaaacg 60
 QY 574 aaagcaaaagcagtgatagacaaattttgttaaaagctagagaagaaattacaaattcgaa 633
 |||||
 Db 61 aaagcaaaagcagtgatagacaaattttgttaaaagctagagaagaaattacaaattcgaa 120
 QY 634 tgcggaattg 643
 |||||
 Db 121 tgcggaattg 130

RESULT 7

AAI55243
 ID AAI55243 standard; DNA; 130 BP.

XX AC AAI55243;
 XX DT 17-OCT-2001 (first entry)
 DE Probe #23929 used to measure gene expression in human placenta sample.
 DE Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.

XX OS Homo sapiens.
 XX PN WO200157272-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00663.
 XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488997/53.
 XX DR Human genome-derived single exon nucleic acid probes useful for
 XX PT analyzing gene expression in human placenta -
 XX PS Claim 25; SEQ ID No 23929; 654pp; English.

XX CC The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.

XX SQ Sequence 130 BP; 60 A; 18 C; 24 G; 28 T; 0 other;
 Query Match 5.5%; Score 130; DB 22; Length 130;
 Best Local Similarity 100.0%; Pred. No. 3.4e-19;

CC and for treating ovarian cancer in a patient. An antigen presenting cell
 CC that expresses the sequences is useful for treating ovarian cancer by
 CC incubating CD4+ and/or CD8+ T cells isolated from a patient. The T cells
 CC can then be proliferated and administered to the patient to inhibit the
 CC development of cancer. The DNA sequences are useful as probes or primers
 CC for nucleic acid hybridisation, to direct expression of a polypeptide in
 CC appropriate host cells. Detecting the presence of a cancer in a patient
 CC involves obtaining a biological sample from the patient, contacting the
 CC biological sample with an agent that binds to the protein, detecting the
 CC amount of protein that binds to the agent, comparing the amount of
 CC protein to a predetermined cut-off value and determining the presence of
 CC cancer. Sequences ABK09464-ABK09802 represent PCR primers and cDNA
 CC molecules encoding ovarian tumour proteins of the invention.
 XX
 SQ Sequence 102 BP; 24 A; 29 C; 28 G; 20 T; 1 other;

Query Match 2.0%; Score 47; DB 24; Length 102;
 Best Local Similarity 100.0%; Pred. No. 0.47;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttgtaccgagctcggtacccctagtaacggcgcgcagtgctgga 47
 |||||
 Db 9 ttgtaccgagctcggtacccctagtaacggcgcgcagtgctgga 55

RESULT 14
 AAS49363/c
 ID AAS49363 standard; DNA; 197 BP.
 XX
 AC AAS49363;
 DT 13-FEB-2002 (first entry)
 XX
 DE Staphylococcus aureus cellular proliferation inhibitory sequence #587.

XX
 KW Antisense; ss; prokaryotic cellular proliferation;
 KW antibiotic; antibacterial; drug design.

OS Staphylococcus aureus.

PN W0200170955-A2.

XX 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

XX 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253623P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX WPI; 2001-611495/70.

XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids .
 XX
 PS Claim 1; Seq ID No 1940; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The

CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence is an antisense
 CC oligonucleotide of the invention.

CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 197 BP; 53 A; 40 C; 34 G; 70 T; 0 other;

Query Match 2.0%; Score 46.8; DB 23; Length 197;
 Best Local Similarity 57.5%; Pred. No. 0.62;
 Matches 84; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 1279 caaatagagagcttaaaaaggtgtacatatcctaatgtccactcccggaagtgaat 1338
 |||||
 Db 196 CAAATTAAGCCCTGTGAAAAAGGCCCAAAATCGTAGTGGACACCTGGCGGTATC 137

QY 1339 gatctgcaaatgagtaactctgcaatctgaagaataataacactactgtgttttagatgaa 1398
 |||||
 Db 136 GACCATTTAAATCGTCGACATTAATAAGCGAGCAATTCATACATTTGATTTAGATGAA 77

QY 1399 gcagacaagatgttgacatgggatt 1424
 |||||

Db 76 GCTGATGAATGATGATATGGGATT 51

RESULT 15

AAH71353/c

ID AAH71353 standard; cDNA; 247 BP.

XX AAH71353;

XX 19-SEP-2001 (first entry)

XX Human cervical cancer marker nucleic acid 2627.

DE Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX Homo sapiens.

XX W0200142467-A2.

XX 14-JUN-2001.

XX 08-DEC-2000; 2000WO-US33312.

XX 08-DEC-1999; 99US-0169681.

PR 21-DEC-1999; 99US-0171350.

PR 14-MAR-2000; 2000US-0189315.

PR 12-MAY-2000; 2000US-0203791.

PR 09-JUN-2000; 2000US-0210600.

XX 21-JUL-2000; 2000US-0220114.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Deeds J, Berger A, Zhao X;

XX WPI; 2001-375006/39.

XX New isolated nucleic acid for diagnosing and treating cervical cancer
 PT and for assessing and detecting compounds for treating the cancer -

XX Claim 1; Page 544; 1051pp; English.

XX The invention relates to novel genes (AAH68727-AAH73383) associated with

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 19:15:13 ; Search time 2465.76 Seconds
(without alignments)
12945.420 Million cell updates/sec

Title: US-09-923-831-42

Perfect score: 2365

Sequence: 1 ttggtacagactcgatcc.....aaaaaaaaaaaaaaaaaaaaa 2365

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 5562878

Minimum DB seq length: 0

Maximum DB seq length: 300

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estbm:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pin:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	243.8	10.3	247	9	BE144935 CM2-HT018
C	210.4	8.9	231	10	BE377169 PM2-TN010
C	154.6	6.5	289	9	AA932998 on91f04.s
4	136.4	5.8	280	10	BC066058
5	115.6	4.9	289	9	BE068292
6	103.8	4.4	270	10	BF153942
7	96.8	4.1	255	9	AA319350
8	95.2	4.0	259	9	AA337335
9	95	4.0	257	9	AA301156
10	94	4.0	293	10	BI507376
11	92.8	3.9	290	9	AA370389
12	92.4	3.9	248	9	AV256222
13	91.2	3.9	254	10	BG688809
C	76.6	3.2	287	9	AA390434
14	74	3.1	200	9	AA054227
15	74	3.1	244	10	N31074
16	74	3.1	244	10	N31074
17	71.6	3.0	276	9	AW114510

18	70.2	3.0	300	9	BB161197
19	69.2	2.9	223	10	BF598412
C	68.4	2.9	277	9	AI641123
20	68.4	2.9	277	9	AI641123
21	66.2	2.8	211	10	H49812
22	65.4	2.8	182	10	N44240
23	65.4	2.8	289	9	AA944137
24	65.2	2.8	263	10	BG635374
25	65	2.7	294	9	AV408925
26	64.8	2.7	243	9	AA750925
27	64.8	2.7	278	9	BE120001
28	64.6	2.7	263	10	BF609454
C	64.2	2.7	261	9	AM491036
29	63.4	2.7	249	10	BE528509
30	63.4	2.7	225	10	Z17903
31	63.2	2.7	256	9	BE029873
32	62.8	2.7	205	9	AV348236
33	62.8	2.6	283	9	AU173638
34	62.2	2.6	219	10	W34581
C	62	2.6	150	9	AA889106
35	61.6	2.6	188	10	BM110125
36	61.6	2.6	188	10	BM110125
C	60.8	2.6	234	9	AV589058
37	60.8	2.6	200	10	BF376705
38	60.6	2.6	256	10	N65778
39	60.6	2.6	280	9	BE597009
40	60.6	2.6	218	9	AA356844
41	60.4	2.6	293	10	BF010681
42	60.4	2.5	251	9	AA363887
43	60.2	2.5	193	9	AA373151
44	60	2.5	193	9	AA373151
45	59.6	2.5	193	9	AA373151

ALIGNMENTS

RESULT

BE144935

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

247 bp mRNA linear
CM2-HT0184-061099-018-g11 HT0184 Homo sapiens cDNA, mRNA sequence.
BE144935
CM2-HT0184-061099-018-g11 HT0184 Homo sapiens cDNA, mRNA sequence.
BE144935
BE144935.1 GI:8607659
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 247)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, P.F., Goldman, G.H., Carvalho, A.E., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=at2-cM2-HT0184-061099-018-g11&ts=1999-10-06&tl=1)
Seq primer: puc 18 forward
High quality sequence stop: 247.
Location/Qualifiers
1. .247
/organism="Homo sapiens"
/db_xref="taxon:9606"

Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

source

1. 289
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="RIKEN full-length enriched, 15 days embryo male testis"
/sex="male"
/tissue_type="testis"
/dev_stage="15 days embryo"
/lab_host="DH10B"
/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research, in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCAGAGCTCTTTTATTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Not = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATTCGAGTAAATTAATTAATCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda PLC I. Cloning sites, 5' end: SalI; 3' end: BamHI".

BASE COUNT
ORIGIN

105 a 44 c 60 g 80 t

Query Match 4.9%; Score 115.6; DB 9; Length 289;
Best Local Similarity 67.7%; Pred. No. 5.6e-08;
Matches 199; Conservative 0; Mismatches 79; Indels 16; Gaps 2;

QY 2029 agacgaatcagagattccagagagcgttgatcaatggctgagaggtttgagcacat 2088
Db 1 AGACCACATCAGCATTTCCCGAGTGGTACTCTCTTAATGCCTAAGAGATACAAACAAAT 60

QY 2089 caacggaaagggaatggaaagaaatggaaagaccccaaggagcccaaggattt 2148
Db 61 AATACAAAGAGAGATGGAAAAAATGGGAAGACCTCAAGAAAGCCCGCAGAGTTT 120

QY 2149 cattaatgtctctgtactagtggtggtgagagaattccaagatttttttagaaatatagtaag 2208
Db 121 TACTAATGTTGGCTGGAGCAGCTAGTGCOCGAGGATTCAGAGTTGTAGTCAAGTG--- 176

QY 2209 scagaagtattggacattttggcagtagtgaagagacggcagctgattgactatttaa 2268
Db 177 ---GAAGTGTGAACATACATCTACTATGGAGAGAC-----TTGGCTGATTCATAA 224

QY 2269 aataatagtgttgaaatatagataccagtggttttttacttttttacttttttaataaaa 2322
Db 225 AATAATAGTGTGTAATAGTAGAATTCATGTTTATACATTTCTTTTATATAAA 278

RESULT 6
BF153942
LOCUS 270 bp mRNA linear EST 01-DEC-2000
DEFINITION 022D01 Mature tuber lambda ZAP Solanum tuberosum cDNA 5' similar to RNA HELICASE-LIKE PROTEIN DB10 gi|1084413|. . . sp|P46942, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BF153942
BF153942.1 GI:11035882
EST
Potato.

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

1 (bases 1 to 270)

Nielsen,K.L, Crookshanks,M., Emmersen,J. and Welinder,K.G.

EST-sequencing of mature potato tuber (Var. Kuras)

Unpublished (2000)

Contact: Karen G. Welinder

Institut for Biotechnology

Aalborg Universitet

Sohngaardsholmsvej 49, 9000 Aalborg, Denmark

Tel: +45 96358467

Fax: +45 98141808

Email: kwebio.auc.dk

High quality sequence stop: 270

POLVA-No.

Location/Qualifiers

1. 270

/organism="Solanum tuberosum"

/cultivar="Field grown Kuras"

/db_xref="taxon:4113"

/clone_lib="Mature tuber lambda ZAP"

/tissue_type="tuber"

/note="Vector: Lambda ZAP"

BASE COUNT 73 a 45 c 80 g 72 t

ORIGIN

Query Match 4.4%; Score 103.8; DB 10; Length 270;

Best Local Similarity 63.3%; Pred. No. 3.3e-06;

Matches 159; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1249 tgggtatattggtggaaatagatgaacaaatagaagagcttaaaaaaggtgtagat 1308

Db 11 TGCTGTATGAGGTGCTCAAAAGGTCTCACTCCGAGAGCTGAGTAGGGGTAGAT 70

QY 1309 atcataatgcaactcccggaagtgaatgatctgcaaatgagttaacttcgtaactcg 1368

Db 71 ATTGTTGTGCGACATCTCTGCTGTTTGAATGACATATTGGAATGAGAGGCTTAGTCTT 130

QY 1369 aagaatataaactactctgtgttttagatgaagcaagaatgttgacatgggatttgaa 1428

Db 131 GGTCAAGTTTCTTAGCTGGTGTAGATGAAGCAGATCGTATGCTAGACATGGGATTGAA 190

QY 1429 cccagataatgaagattttgttagatgtgcccagatagggcagacagttatgaccagt 1488

Db 191 CCTCAGATGAAGAAGATTGGAAGAGGTGGCTACCGAGGCGAGCATTGATGTATACA 250

QY 1489 gtcacatggcc 1499

Db 251 GATACATGGCC 261

RESULT 7

AA319350

LOCUS

DEFINITION

AA319350

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

human.

Human sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 255)

REFERENCE

AUTHORS

Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult

C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White

O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, J.M., Fitzhugh, W.M., Fritchman, J.L., Geohagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Springs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferris, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl.), 3-174 (1995)

96026280

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi.html>)

Seq primer: M13 Reverse.

FEATURES

source

1. .255

/organism="Homo sapiens"

/db_xref="ATCC (inhost):119896"

/db_xref="taxon:9606"

/clone_lib="Adrenal gland tumor"

/dev_stage="adult"

/note="Organ: adrenal gland; Vector: pBluescript SK-"

Site_1: EcoRI; Site_2: XhoI

80 a 42 c 63 g 56 t 4 others

BASE COUNT

ORIGIN

Query Match 4.1%; Score 96.8; DB 9; Length 255;

Best Local Similarity 65.3%; Pred. No. 3.7e-05;

Matches 154; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

QY 1248 ttgtgtatatgtgtgtggaatagagatgaacaaatagagagcttaaaaaaggtgtaga 1307

DB 17 TTGTATCTACGGTGTGCTCTTAAGGGACCACAAATACGTNATTTGGAGAGGTTGGA 76

QY 1308 tatcataattgaactcccggaagattgaatctgcataatgagtaacttcgtcaatct 1367

DB 77 AATCTGTATTGCAACACCTGGAAGACTGATTGTTTNNAGAGTGTGGAACCAATCT 136

QY 1368 gaagaataaactactctgttttagatgaagcagacagatgttgacatgggattga 1427

DB 137 NAGAAGACACACCTACCTTGTCTTGAAGACAGATAGATGCTTGNATTTGGCTTGA 196

QY 1428 acccagataaagagatttggtta-gatgtgcgcgcagatagcagacagttag 1482

DB 197 ACCCAATAAGGAAGATTGTGGATCAATAAAGACCTGATAGGCAAACTTAATG 252

RESULT

8

AA337335

LOCUS

DEFINITION

AA337335

VERSION

KEYWORDS

SOURCE

human.

259 bp mRNA linear EST 21-APR-1997

EST42053 Endometrial tumor Homo sapiens cDNA 5' end similar to similar to RNA helicase p68, mRNA sequence.

AA337335.1 GI:1989573

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 259)

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Iee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, J.M., Fitzhugh, W.M., Fritchman, J.L., Geohagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Springs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferris, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

TITLE

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6547 Suppl.), 3-174 (1995)

96026280

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

JOURNAL

MEDLINE

COMMENT

FEATURES

source

1. .259

/organism="Homo sapiens"

/db_xref="ATCC (inhost):139075"

/db_xref="taxon:9606"

/clone_lib="Endometrial tumor"

/sex="female"

/dev_stage="adult"

/note="Organ: endometrium; Vector: pBluescript SK-"

Site_1: EcoRI; Site_2: XhoI

80 a 42 c 66 g 68 t 3 others

BASE COUNT

ORIGIN

Query Match 4.0%; Score 95.2; DB 9; Length 259;

Best Local Similarity 65.1%; Pred. NO. 6.5e-05;

Matches 153; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

QY 1248 ttgtgtatatgtgtgtggaatagagatgaacaaatagagagcttaaaaaaggtgtaga 1307

DB 17 TTGTATCTACGGTGTGCTCTTAAGGGACCACAAATACGTGATTGGAGAGGTTGGA 76

QY 1308 tatcataattgaactcccggaagattgaatctgcataatgagtaacttcgtcaatct 1367

DB 77 AATCTGTATTGCAACACCTGGAAGACTGATTGANTTTTATAGTGTGGAAAACCAATCT 136

QY 1368 gaagaataaactactctgttttagatgaagcagacagatgttgacatgggattga 1427

DB 137 GAGAAGANCAACCTACCTGTCTCTTGTATGACAGATAGATGCTTGNATTTGGCTTGA 196

QY 1428 acccagataaagagattt-gatgtgcgcgcagatagcagacagttag 1481

DB 197 ACCCAATAAGGAAGATTGTGGATCAATAAAGACCTGATAGGCAAACTTAAT 251

RESULT

9

AA301156


```

Db 20 GAAGCACTGTATTGTTGGTGGTCGCCAAAGAGAACCCAGACTCGTGTATTAGACAAAG 79
Qy 1301 gtgtgatatacaattgaactcccggaagattgaatgatctgcacaaatgagtaacttgg 1360
Db 80 GAGTTGAGATTGTCATTCCTACCTGCGTGGACGATTAATGATTGTTTGGAAAGCTGGCACAA 139
Qy 1361 tcaattgaagaataataaccactctggttttagatgaagcagacagatgttgacatgg 1420
Db 140 CTAATTACCTAGATGTCACATATTGGTATTGGATGAGCAGATGATGTTAGATAGG 199
Qy 1421 gattgaacccacagataatgaagattttgttagatgtgcccacagatgagcagacagttta 1480
Db 200 GTTTTGAACACAGATTAGAAAATATTGTAACAAATCAGACAGATGACACAGATTAA 259
Qy 1481 tgaccagtgctacatggccctcattcagttc 1510
Db 260 TGTGTCGTGCACTGTGCCAAAGAGAGTTC 289

RESULT 11
AA370389 290 bp mRNA linear EST 21-APR-1997
LOCUS EST02007 prostate gland I Homo sapiens cDNA 5' end similar to
DEFINITION similar to RNA helicase p68, mRNA sequence.
ACCESSION AA370389
VERSION AA370389.1 GI:2022706
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 290)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Feldner,R.A., Bult
,C.O., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
,O., Sutton,G., Blake,J.A., Brandon,K.C., Man-Whi,C., Clayton,R.A.,
Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald
,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S., Glodek,A.,
Gneltm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M.,
Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palaoques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.P., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,I.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W.,
Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L.,
Kunsch,C., Rungtun,J., Li,H., Meissner,P.S., Olsen,H., Raymond,L.,
Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon
,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M. and
Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl.), 3-174 (1995)
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tldb/hgi/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. 290
/organism="Homo sapiens"
/db_xref="atcc (inhost):174856"
/db_xref="taxon:9606"
/clone_lib="prostate gland I"
/sex="male"
/dev_stage="adult, 21 yrs"
/note="Organ: prostate; Vector: pBluescript SK-; Site:1."

```

```

BASE COUNT 89 a 52 c 75 g 74 t
ORIGIN
Query Match 3.9%; Score 92.8; DB 9; Length 290;
Best Local Similarity 66.5%; Pred. No. 0.00014;
Matches 133; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
Qy 1248 tigtgtatattgggtgggaataagatgagataagagagcttaaaaaagttgtaga 1307
Db 71 TTGTATACCTGGTGGTCTCTAAGGACCAAAATACGCTATTGGAGAGAGGTGGA 130
Qy 1308 tatcataattcaactcccggaagattgaatgctgcacaaatgagtaactctcact 1367
Db 131 ATCTGTATTGTCACACCTGGAGACATGATTGACTTTTAGAGTGTGAAAACCAATCT 190
Qy 1368 gaagaataacactactgttttagatgaagcagacagatgttgacatgggattga 1427
Db 191 GAGAAGAACAACCTACCTTGTCTCTGTAGTGAAGCAGATAGATGCTTGATATGGCTTGA 250
Qy 1428 accccagataataagattt 1447
Db 251 ACCCCAAATAGGAGGTTT 270

RESULT 12
AV256222 248 bp mRNA linear EST 04-NOV-1999
LOCUS AV256222 RIKEN full-length enriched, adult male testis (DH10B) Mus
DEFINITION musculus cDNA clone 492156D21 3', mRNA sequence.
ACCESSION AV256222.1 GI:6243681
VERSION AV256222.1
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 248)
Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,
Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai
,C., Kawai,O., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,
Masuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata
,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H.,
Suzuki,H., Takahashi,F., Tateo,M., Tomimaga,N., Tsunoda,Y.,
Wataniki,A., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T.,
Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Konno,H., et al. 1999)
Unpublished (1999)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Sasaki,N., Izawa,M., Wataniki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,
Matsura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki
,Y.
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)
Itoh,M., Kitsuai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
,Y. and Hayashizaki,Y.
Automated filtration-based high-throughput plasmid preparation
system. Genome Res. 9 (5), 463-470 (1999)
Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)

```

RESULT	13
EG688809	
LOCUS	254 bp mRNA linear EST 02-MAY-2001
DEFINITION	33579 BARC 50V Bos taurus cDNA 5', mRNA sequence.
ACCESSION	EG688809
VERSION	EG688809.1 GI:13930610
KEYWORDS	EST.
SOURCE	COW.
ORGANISM	Bos taurus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
	Bovidae; Bovinae; Bos.
	1 (bases 1 to 254)
REFERENCE	Sonstegard,T.S.; Capuco,A.V.; Van Tassell,C.P.; Ashwell,M.S. and
AUTHORS	

RESULT	14
AW390434/c	
LOCUS	
DEFINITION	287 bp mRNA linear EST 04-FEB-2000
ACCESSION	RCO-ST0185-231099-032-f03 ST0185 Homo sapiens cDNA, mRNA sequence.
VERSION	AW390434
KEYWORDS	AW390434.1 GI:6895093
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	1 (bases 1 to 287)
	HCGP http://www.ludwig.org.br/ORSTES .
	The FAPESP/LICR Human Cancer Genome Project
	Unpublished (1999)

GenCore version 4.5
 Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 19:12:38 ; Search time 4252.24 Seconds
 (without alignments)
 11638.868 Million cell updates/sec

Title: US-09-923-831-42
 Perfect score: 2365
 Sequence: 1 ttggtagcagctcgatcc.....aaaaaaaaaaaaaa 2365

Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
 Total number of hits satisfying chosen parameters: 843946

Minimum DB seq length: 0
 Maximum DB seq length: 100

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.inu.*

20: em.om.*

21: em.or.*

22: em.ov.*

23: em.pat.*

24: em.ph.*

25: em.pl.*

26: em.ro.*

27: em.sts.*

28: em.un.*

29: em.vi.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htgo.inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description

1	39.4	1.7	99	6	A08898	A08898 H.sapiens (
2	39.2	1.7	86	6	AX203294	AX203294 Sequence
3	37	1.6	98	12	SYNROCBA	M63418 Synthetic 5
4	36.8	1.6	84	6	AR038854	AR038854 Sequence
5	36.8	1.6	84	6	AR050387	AR050387 Sequence
6	36.4	1.5	80	6	AR011130	AR011130 Sequence
7	36.2	1.5	74	6	A08894	A08894 H.sapiens (
8	36	1.5	77	6	A08908	A08908 H.sapiens (
9	36	1.5	96	6	A08909	A08909 H.sapiens (
10	35.8	1.5	85	6	BD003061	BD003061 Polynucle
11	35.6	1.5	73	6	A08915	A08915 H.sapiens (
12	35.4	1.5	85	6	AX099459	AX099459 Sequence
13	35.4	1.5	97	6	AX260590	AX260590 Sequence
14	35.4	1.5	98	6	AX320854	AX320854 Sequence
15	35	1.5	100	9	AF117819	AF117819 Homo sapi
16	34.8	1.5	62	6	AX205289	AX205289 Sequence
17	34.8	1.5	100	6	AX260734	AX260734 Sequence
18	34.6	1.5	80	6	A08907	A08907 H.sapiens (
19	34.6	1.5	80	6	A18777	A18777 5' terminus
20	34.6	1.5	80	6	AR001312	AR001312 Sequence
21	34.6	1.5	80	6	AR130410	AR130410 Sequence
22	34.6	1.5	84	6	A08920	A08920 H.sapiens (
23	34.6	1.5	85	6	AX182181	AX182181 Sequence
24	34.6	1.5	91	3	DI8ACTIN8A	M25215 Dictyosteli
25	34.4	1.5	71	9	S76508	S76508 D188 (B) (h
26	34.4	1.5	79	6	AX099452	AX099452 Sequence
27	34.4	1.5	89	6	AX203299	AX203299 Sequence
28	34.2	1.4	86	6	A08895	A08895 H.sapiens (
29	34.2	1.4	90	6	AR148129	AR148129 Sequence
30	34.2	1.4	90	6	EL2580	EL2580 Probe, 6/19
31	34	1.4	86	3	DDI8CTC	M29111 D.discoideu
32	33.8	1.4	93	6	AX320853	AX320853 Sequence
33	33.8	1.4	98	6	AX209704	AX209704 Sequence
34	33.8	1.4	100	10	RN012531	G12531 Rattus norv
35	33.6	1.4	76	6	AX260960	AX260960 Sequence
36	33.6	1.4	79	6	AX182145	AX182145 Sequence
37	33.6	1.4	81	6	AX182108	AX182108 Sequence
38	33.2	1.4	62	6	AX205290	AX205290 Sequence
39	33.2	1.4	62	6	AX205291	AX205291 Sequence
40	33.2	1.4	70	6	AX260894	AX260894 Sequence
41	33	1.4	69	6	BD005374	BD005374 Secreted
42	33	1.4	69	6	BD005384	BD005384 Secreted
43	33	1.4	69	6	I89934	I89934 Sequence 7
44	33	1.4	69	6	I89944	I89944 Sequence 22
45	33	1.4	74	6	AX322391	AX322391 Sequence

ALIGNMENTS

RESULT	1	A08898	99 bp	DNA	linear	PAR 02-SEP-1993
LOCUS						
DEFINITION		H.sapiens (haplotype 2B, allele MS32, isolate French, serial number 16) minisatellite sequence.				
ACCESSION		A08898				
VERSION		A08898.1				
KEYWORDS		GI:411820				
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS		Jeffreys,A.J.				
TITLE		Extended nucleotide sequences				
JOURNAL		Patent: EP 0370719-A 81 30-MAY-1990;				
FEATURES		IMPERIAL CHEMICAL INDUSTRIES PLC				
SOURCE		Location/Qualifiers				
		1. .99				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
BASE COUNT		67 a	0 c	0 g	32 t	
ORIGIN						

```

FEATURES
  LOCATION_QUALIFIERS
    1.98
    /organism="synthetic construct"

```

Query Match	1.6%	Score 36.8;	DB 6;	Length 84;
Best Local Similarity	77.2%	Pred. No. 1.2e+03;		

Matches 44; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2309 ttctttaaataaataagagctatttaaaacttaaaaaaataaaaaaataaaaaa 2365
 Db 1 TTTTNTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 57

RESULT 6
 LOCUS AR011130 80 bp DNA linear PAT 04-DEC-1998
 DEFINITION Sequence 6 from patent US 5762924.
 ACCESSION AR011130
 VERSION AR011130.1 GI:3969120
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 80)
 AUTHORS Dall,D.James., Fernon,C.Anne. and Sriskantha,A.
 TITLE Recombinant entomopoxvirus
 JOURNAL Patent: US 5762924-A 6 09-JUN-1998;
 FEATURES Location/Qualifiers
 source 1..80
 BASE COUNT 46 a 8 c 2 g 24 t
 ORIGIN

Query Match 1.5%; Score 36.4; DB 6; Length 80;
 Best Local Similarity 70.0%; Pred. No. 1.5e+03;
 Matches 45; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 2295 ccagtggtttatactttctttaaataaataagagctatttaaaacttaaaaaaataaaaaa 2354
 Db 5 CCAATATTTTACAACTCTAATAAAAAATAGATAAATTTATTATTATAAATAGCAAA 64

QY 2355 aaaaaaataa 2364
 Db 65 AAATAAAAAA 74

RESULT 7
 LOCUS A08894 74 bp DNA linear PAT 02-SEP-1993
 DEFINITION H.sapiens (haplotype 2A, allele MS32, isolate English, serial number 9) minisatellite sequence.
 ACCESSION A08894
 VERSION A08894.1 GI:411816
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 74)
 AUTHORS Jeffreys,A.J.
 TITLE Extended nucleotide sequences
 JOURNAL Patent: EP 0370719-A 77 30-MAY-1990;
 IMPERIAL CHEMICAL INDUSTRIES PLC
 FEATURES Location/Qualifiers
 source 1..74
 BASE COUNT 52 a 0 c 0 g 22 t
 ORIGIN

Query Match 1.5%; Score 36.2; DB 6; Length 74;
 Best Local Similarity 72.3%; Pred. No. 1.6e+03;
 Matches 47; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 2301 ttttatactttctttaaataaataagagctatttaaaacttaaaaaaataaaaaaataaaaaa 2360
 Db 2 TTTTATATTTTATTAATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 61

QY 2361 aaaaa 2365
 Db 62 AAAAA 66

RESULT 8
 LOCUS A08908 77 bp DNA linear PAT 02-SEP-1993
 DEFINITION H.sapiens (haplotype 3, allele MS32, isolate Mormon, serial number 10) minisatellite sequence.
 ACCESSION A08908
 VERSION A08908.1 GI:411830
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 77)
 AUTHORS Jeffreys,A.J.
 TITLE Extended nucleotide sequences
 JOURNAL Patent: EP 0370719-A 91 30-MAY-1990;
 IMPERIAL CHEMICAL INDUSTRIES PLC
 FEATURES Location/Qualifiers
 source 1..77
 BASE COUNT 72 a 0 c 0 g 5 t
 ORIGIN

Query Match 1.5%; Score 36; DB 6; Length 77;
 Best Local Similarity 80.8%; Pred. No. 1.8e+03;
 Matches 42; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2314 ttaataaataaataagagctatttaaaacttaaaaaaataaaaaaataaaaaaataaaaaa 2365
 Db 1 TTATATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 52

RESULT 9
 LOCUS A08909 96 bp DNA linear PAT 02-SEP-1993
 DEFINITION H.sapiens (haplotype 3, allele MS32, isolate English, serial number 15) minisatellite sequence.
 ACCESSION A08909
 VERSION A08909.1 GI:411831
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 REFERENCE 1 (bases 1 to 96)
 AUTHORS Jeffreys,A.J.
 TITLE Extended nucleotide sequences
 JOURNAL Patent: EP 0370719-A 92 30-MAY-1990;
 IMPERIAL CHEMICAL INDUSTRIES PLC
 FEATURES Location/Qualifiers
 source 1..96
 BASE COUNT 91 a 0 c 0 g 5 t
 ORIGIN

Query Match 1.5%; Score 36; DB 6; Length 96;
 Best Local Similarity 80.8%; Pred. No. 1.8e+03;
 Matches 42; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2314 ttaataaataaataagagctatttaaaacttaaaaaaataaaaaaataaaaaaataaaaaa 2365
 Db 1 TTAATAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 52

aps	0;
-APR-2001	
ostomi;	
Evans,C.,	
Ino,M.J.	
aps	0;
2346	
66	
-OCT-2001	
ostomi;	
n cancer	

```
1. .100
/organism="Homo sapiens"
/db xref="taxon:9606"
```

Search completed: September 24, 2002, 21:07:56
Job time: 6918 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 19:14:08 ; Search time 370.16 Seconds
(without alignments)
10969.588 Million cell updates/sec

Title: US-09-923-831-42

Perfect score: 2365

Sequence: 1 ttgtaccgagctcgatcc.....aaaaaaaaaaaaaaaaaaaaa 2365

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 2046006

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq_032802:*

1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39.2	1.7	86	22	AA523317 Human prostate can
2	38.6	1.6	95	13	AAQ28050 Entomopoxvirus sph
3	37.8	1.6	79	19	AAV28706 Human novel secret
4	36.8	1.6	84	19	AAV09269 Nucleotide sequenc
5	36.8	1.6	84	19	AAV04275 Secreted protein C
c 6	36.6	1.5	51	22	AA132186 Human SNP oligonuc
c 7	36.6	1.5	99	18	AAT91300 Human M97-2 secret
c 8	36.4	1.5	92	18	AAT88439 Padlock probe alph
c 9	36.2	1.5	90	21	AAF18395 Lung cancer associ

c 10	36	1.5	53	22	ABA11098 Human nervous syst
c 11	36	1.5	90	12	AA011762 Self-complementary
c 12	36	1.5	94	16	AA089696 Vector back-end ol
c 13	35.8	1.5	80	19	AAV37197 Oligonucleotide se
c 14	35.8	1.5	85	19	AAV05720 Nucleotide sequenc
c 15	35.6	1.5	51	22	AA129620 Human SNP oligonuc
c 16	35.4	1.5	85	19	AAV32418 Homo sapiens clone
c 17	35.4	1.5	85	22	AAV98434 Human cDNA clone B
c 18	35.4	1.5	97	23	AA57565 cDNA #241 encoding
c 19	35.4	1.5	98	12	AA011764 Self-complementary
c 20	35	1.5	93	21	AA098487 Human colon cancer
c 21	35	1.5	94	16	AA089695 Vector back-end ol
c 22	34.8	1.5	61	21	AA024838 Human secreted pro
c 23	34.8	1.5	62	22	AA013991 Deletion construct
c 24	34.8	1.5	100	23	AA557709 cDNA #385 encoding
c 25	34.6	1.5	63	21	AA014751 Human secreted pro
c 26	34.6	1.5	81	21	AA013249 Human secreted pro
c 27	34.6	1.5	85	22	AA07817 Cervical cancer pr
c 28	34.4	1.5	79	19	AAV32414 Homo sapiens clone
c 29	34.4	1.5	79	22	AA098430 Human cDNA clone B
c 30	34.4	1.5	86	12	AA011760 Self-complementary
c 31	34.4	1.5	89	22	AA523322 Human prostate can
c 32	34.4	1.5	98	20	AA523322 Porcine reproducti
c 33	34.2	1.4	90	18	AA000174 Synthetic DNA prob
c 34	34.2	1.4	90	19	AAV37789 Analytical solid p
c 35	34	1.4	51	22	AA029656 Human SNP oligonuc
c 36	34	1.4	51	22	AA032448 Human SNP oligonuc
c 37	34	1.4	77	22	AA529173 Genomic sequence #
c 38	34	1.4	77	22	AA078766 Human immune/haema
c 39	34	1.4	77	22	AA078766 Human immune/haema
c 40	34	1.4	77	22	AA078766 Human immune/haema
c 41	33.8	1.4	69	18	AA088081 3' portion of cDNA
c 42	33.8	1.4	69	18	AA088081 Human secreted pro
c 43	33.8	1.4	98	22	AA525363 Human ovarian PCR
c 44	33.6	1.4	60	21	AA015704 Human secreted pro
c 45	33.6	1.4	76	23	AA557935 cDNA #611 encoding

ALIGNMENTS

RESULT 1
AA523317
ID AA523317 standard; cDNA; 86 BP.
XX AC AA523317;
XX DT 24-OCT-2001 (first entry)
XX DE Human prostate cancer marker #39.
XX KW Prostate cancer; human; cancer marker; metastasis; diagnostic;
XX KW therapeutic; immunogenic; ss.
XX OS Homo sapiens.
XX PN WO200153836-A2.
XX PD 26-JUL-2001.
XX PF 24-JAN-2001; 2001WO-US023318.
XX PR 24-JAN-2000; 2000US-0178525.
XX PR 17-FEB-2000; 2000US-0183245.
XX PR 16-MAR-2000; 2000US-0190139.
XX PR 31-MAY-2000; 2000US-0208126.
XX PR 18-JUL-2000; 2000US-0219705.
XX PR 13-DEC-2000; 2000US-0255160.
(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege W, Monahan JE;

PT or suppressor, etc.
 PS Disclosure; Page 94; 117pp; English.
 XX
 CC This sequence encodes the 3'-end of a novel secreted protein from clone
 CC BF290.li isolated from a human fetal brain cDNA library. Such secreted
 CC proteins can be used for, e.g. research purposes (as markers for
 CC tissues, molecular weight markers for gels, primers and probes), for
 CC nutrition (as C, N or carbohydrate source), as a cytokine for cell
 CC proliferation and differentiation activity, as immune stimulants or
 CC suppressors, e.g. for viral, bacterial or fungal infections, for
 CC autoimmune diseases such as multiple sclerosis or systemic lupus
 CC erythematosus, to regulate haematopoiesis, for tissue growth, as an
 CC activin or inhibin, or having chemotactic or chemo-knetic, haemostatic
 CC and thrombocytic, receptor/ligand, anti-inflammatory or tumour inhibitory
 CC activities.
 XX
 SQ Sequence 79 BP; 73 A; 0 C; 0 G; 6 T; 0 other;

Query Match 1.6%; Score 37.8; DB 19; Length 79;
 Best Local Similarity 78.9%; Pred. No. 45;
 Matches 45; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 2309 ttctttaaataaagaattatttaaaacttaaaataaaataaaataaaataaa 2365
 ||| |||| |||| ||| ||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 1 ttatttaaaaaataaaataaaataaaataaaataaaataaaataaaataaa 57

RESULT 4
 AAV09269
 ID AAV09269 standard; cDNA; 84 BP.
 XX
 AC AAV09269;
 XX
 DT 07-JUL-1998 (first entry)
 XX
 DE Nucleotide sequence of the 3' portion of the C195_1 protein.
 XX
 KW C195_1 protein; human PBMC cDNA library; cytokine activity; ss;
 KW peripheral blood mononuclear cell; nutritional activity; EST;
 KW cell proliferation/differentiation activity; expressed sequence tag.
 XX
 OS Homo sapiens.
 XX
 PN WC9748801-A2.
 XX
 PD 24-DEC-1997.
 XX
 PF 16-JUN-1997; 97WO-US10501.
 XX
 PR 17-JUN-1996; 96US-0664596.
 XX
 PA (GEMV) GENETICS INST INC.
 XX
 PI Bowman M, Evans C, Jacobs K, Lavallie ER, McCoy JM;
 PI Merberg D, Racie LA, Spaulding V, Treacy M;
 XX
 DR WPI; 1998-063142/06.
 XX
 PT Poly-nucleotide(s) and proteins obtained from human PBMC, dendritic
 PT cell, adult brain, foetal brain and adult testes cDNA libraries -
 PT used in research, detection and therapy of, e.g. cytokine and cell
 PT proliferation or differentiation
 XX
 PS Claim 13; Page 48; 78pp; English.
 XX

CC This nucleotide sequence encodes the 3' portion of the C195_1 protein
 CC which was isolated from a human PBMC cDNA library. The products of
 CC the polynucleotides of the invention can be used in research, detection
 CC and therapy, as they may have nutritional activity, cytokine and cell
 CC proliferation/differentiation activity. A search against the Genbank
 CC database demonstrated that this sequence has at least some homology

CC with two ESTs identified as "ygl106.r1 Homo sapiens cDNA clone
 CC 31937 5'" and "ym15f12.r1 Homo sapiens cDNA clone 48025 5'".
 XX
 SQ Sequence 84 BP; 76 A; 0 C; 0 G; 7 T; 1 other;

Query Match 1.6%; Score 36.8; DB 19; Length 84;
 Best Local Similarity 77.2%; Pred. No. 75;
 Matches 44; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 2309 ttctttaaataaagaattatttaaaacttaaaataaaataaaataaaataaa 2365
 ||| |||| |||| ||| ||| |||| |||| |||| |||| |||| |||| |||| ||||
 Db 1 ttattttnaaaaataaaataaaataaaataaaataaaataaaataaaataaa 57

RESULT 5
 AAV04275
 ID AAV04275 standard; cDNA; 84 BP.
 XX
 AC AAV04275;
 XX
 DT 22-JUN-1998 (first entry)
 XX
 DE Secreted protein C195_1 cDNA 3' end.
 XX
 KW Secreted protein; C195_1; peripheral blood mononuclear cell; PBMC;
 KW protein factor; human; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO9748800-A1.
 XX
 PD 24-DEC-1997.
 XX
 PF 16-JUN-1997; 97WO-US10500.
 XX
 PR 25-OCT-1996; 96US-0738367.
 PR 17-JUN-1996; 96US-0664596.
 PR 27-SEP-1996; 96US-0721936.
 XX
 PA (GEMV) GENETICS INST INC.
 XX
 PI Jacobs K, Lavallie ER, McCoy JM, Merberg D, Racie LA;
 PI Spaulding V, Treacy M;
 XX
 DR WPI; 1998-063141/06.
 XX
 PT Nucleic acid encoding secreted protein from human peripheral blood
 PT mononuclear cells - useful, e.g. as immunomodulators, antitumour
 PT agents, promoters of tissue growth, haemostatic and thrombolytic
 PT agents etc.
 XX
 PS Disclosure; Page 36; 52pp; English.
 XX
 CC This cDNA sequence comprises the 3' end, including the polyA tail,
 CC of cDNA clone C195_1 (see AAV04273), which codes for a novel secreted
 CC protein (see AAV41600) of human peripheral blood mononuclear cells
 CC (PBMC). Clone C195_1 was isolated from a PBMC cDNA library using
 CC methods which are selective for cDNAs encoding secreted proteins.
 CC It is deposited as ATCC 98079. C195_1 can be used to express
 CC recombinant secreted proteins; as a tissue/molecular weight marker;
 CC for chromosome identification; to identify possible genetic
 CC disorders; to isolate new related DNA; as a source of PCR primers;
 CC to generate anti-protein or anti-DNA antibodies; in interaction
 CC trap assays to identify sequences that encode interacting proteins
 CC etc. The C195_1 secreted protein may have a number of biological
 CC activities.
 XX
 SQ Sequence 84 BP; 76 A; 0 C; 0 G; 7 T; 1 other;

Query Match 1.6%; Score 36.8; DB 19; Length 84;
 Best Local Similarity 77.2%; Pred. No. 75;

Matches	44;	Conservative	0;	Mismatches	13;	Indels	0;	Gaps	0;
QY	2309	tttttttaataaataagtagtatttaaaacttaaaaaaataaaaaaataaaaaa 2365							
Db	1	tttttttaaaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaa 57							
RESULT 6									
AAL32186/c									
ID	AAL32186 standard; DNA; 51 BP.								
XX	AC								
XX	AAL32186;								
DT	24-JAN-2002	(first entry)							
XX	Human SNP oligonucleotide #5394.								
DE	XX								
XX	Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;								
KW	neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;								
KW	amyloid protein; angiotensin; apoptosis related protein; cadherin;								
KW	cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;								
KW	complement related protein; cytochrome; kinesin; cytokine; interferon;								
KW	interleukin; G-protein coupled receptor; thioesterase; inflammation;								
KW	multifactorial disease; autoimmune disease; infection;								
KW	nervous system disease; ss.								
OS	Homo sapiens.								
XX									
PN	WO200147944*A2.								
PD	05-JUL-2001.								
PF	28-DEC-2000; 2000WO-US35498.								
PR	28-DEC-1999; 9PUS-0173419.								
PR	27-DEC-2000; 2000US-0173419.								
XX	(CURA-) CURAGEN CORP.								
PA	XX								
PI	Shankets RA, Leach M;								
PI	XX								
DR	WPI; 2001-465210/50.								
XX									
PT	Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,								
PT	oncogenes and histones, useful for diagnosing and treating, e.g.								
PT	cancer, autoimmune diseases and infections -								
XX									
XX	Claim 1; Page 2939; 4143pp; English.								
PS									
CC	The present invention relates to oligonucleotides encoding polymorphic								
CC	variants of proteins related to amylases, amyloid proteins, angiotensin,								
CC	apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,								
CC	histones, kinases, colony stimulating factors, complement related								
CC	proteins, cytochromes, kinesins, cytokines, interferons, interleukins,								
CC	G-protein coupled receptors and thioesterases. The present sequence is								
CC	one such oligonucleotide. The oligonucleotides and the peptides encoded								
CC	by them may be used in the prevention, diagnosis and treatment of								
CC	diseases associated with inappropriate expression of the proteins listed								
CC	above. Disorders that may be prevented, diagnosed and/or treated include								
CC	multifactorial diseases with a genetic component, such as autoimmune								
CC	diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,								
CC	systemic lupus erythematosus and Grave's disease), inflammation, cancer								
CC	(e.g. cancers of the bladder, brain, breast, colon and kidney,								
CC	leukaemia), diseases of the nervous system and an infection of pathogenic								
CC	organisms.								
XX									
SQ	Sequence 51 BP; 12 A; 0 C; 2 G; 37 T; 0 other;								
Query Match 1.5%; Score 36.6; DB 22; Length 51;									
Best Local Similarity 82.4%; Pred. No. 72;									
Matches	42;	Conservative	0;	Mismatches	9;	Indels	0;	Gaps	0;

Query Match: 1.5%; Score 36.4; DB 18; Length 92;


```

PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249220.
PR 17-NOV-2000; 2000US-0249221.
PR 17-NOV-2000; 2000US-0249222.
PR 17-NOV-2000; 2000US-0249223.
PR 17-NOV-2000; 2000US-0249224.
PR 17-NOV-2000; 2000US-0249225.
PR 17-NOV-2000; 2000US-0249226.
PR 17-NOV-2000; 2000US-0249227.
PR 17-NOV-2000; 2000US-0249228.
PR 17-NOV-2000; 2000US-0249229.
PR 17-NOV-2000; 2000US-0249230.
PR 17-NOV-2000; 2000US-0249231.
PR 17-NOV-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI P-PSDB; ABB14772.
XX
XX WPI; 2001-541565/60.
XX P-PSDB; ABB14772.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
XX useful for preventing, diagnosing and/or treating nervous system
XX cancers and metastases -
XX
XX Claim 1; SEQ ID NO 105; 1701pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AB11004-ABA21534) and proteins
XX (AB14678-AB18001) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemias;
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX and parasitic infections.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
XX Sequence 53 BP; 10 A; 17 C; 16 G; 10 T; 0 other;
XX
XX Query Match 1.5%; Score 36; DB 22; Length 53;
XX Best Local Similarity 100.0%; Pred. No. 99;
XX Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 12 ctggatccctagtaacgcccagtgctgtgaa 47
XX |||||||||||||||||||||||||||||||
XX Db 10 ctggatccctagtaacgcccagtgctgtgaa 45
XX
XX RESULT 11
XX AAQ11762/c
XX ID AAQ11762 standard; DNA; 90 BP.
XX

```

```

AC AAQ11762;
XX
XX 22-JUL-1991 (first entry)
XX
XX Self-complementary, T3 promoter hairpin-forming sequence #2.
DE
XX T3 bacteriophage; DNA-dependent RNA polymerase;
XX target sequence amplification; ss.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX stem_loop 1..90 a
XX /note= "number of T's in loop is 50"
XX
XX EP427073-A.
XX
XX 15-MAY-1991.
XX
XX 27-OCT-1990; 90EP-0120650.
XX
XX 23-AUG-1990; 90US-0569991.
XX 09-NOV-1989; 89US-0434372.
XX
XX (MOLE-) MOLECULAR DIAGNOSTI.
XX
XX Dattagupta N;
XX
XX WPI; 1991-141638/20.
XX
XX Nucleic acid probe for amplification and detection of target
XX sequence - capable of forming ligatable hairpin structured
XX promoter and transcribing target sequence, is sensitive and
XX useful in medical diagnosis
XX
XX Claim 2; Page 11; 15pp; English.
XX
XX The sequence is an example of a preferred T3 RNA polymerase
XX promoter for use in the invention. There can be from 2 to 50 T's in
XX the loop region. A probe sequence is ligated to the 3' end of the
XX promoter region. Upon hybridisation of the probe to a target sequence
XX and ligation of the hybridised target sequence to the 5' end of the
XX hairpin-forming sequence, the target sequence can be transcribed by
XX T3 RNA polymerase. This allows the target sequence to be amplified
XX (and detected) using a single oligonucleotide component (c.f. PCR).
XX See also AAQ11759-Q11761 and AAQ11763-Q11764.
XX
XX Sequence 90 BP; 12 A; 8 C; 8 G; 62 T; 0 other;
XX
XX Query Match 1.5%; Score 36; DB 12; Length 90;
XX Best Local Similarity 80.8%; Pred. No. 1.1e+02;
XX Matches 42; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
XX 2314 ttaataaaataagaactatttaaaacttaaaaaaaataaaaaaaataaaaaa 2365
XX ||||||||| ||| ||| |||||||||||||||||||
XX Db 75 TTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 24
XX
XX RESULT 12
XX AAQ89696
XX ID AAQ89696 standard; DNA; 94 BP.
XX
XX AAQ89696;
XX
XX 07-NOV-1995 (first entry)
XX
XX Vector back-end oligonucleotide (complementary strand).
XX
XX Universal cloning vector; gene expression; cDNA expression;
XX back-end oligonucleotide; ss.
XX

```

OS Synthetic.
 PN WO9510620-A.
 XX
 PD 20-APR-1995.
 XX
 PF 14-OCT-1994; 94WO-US11719.
 XX
 PR 15-OCT-1993; 93US-0136148.
 XX
 PA (MLCW) MALLINCKRODT VETERINARY INC.
 XX
 PI McMullen JR, Syrenki RM, Zook CA;
 XX
 DR WPI; 1995-161805/21.
 XX
 PT Universal cloning vector for expression of heterologous genes or
 PT cDNA - allows high levels of expression without the need for
 PT sequence modification
 XX
 PS Claim 9; Fig.3; 40pp; English.
 XX
 CC A universal vector comprises a pl promoter, a ribosome binding site,
 CC the first 30 bp of the delta-7 pig somatostatin gene, a 21 bp front-
 CC end oligonucleotide (pref. the dimer comprising the complementary
 CC strands given in AAQ89693-94) containing multiple blunt end restriction
 CC enzyme sites, a back-end oligonucleotide (pref. the dimer
 CC comprising the strands given in AAQ89695-96) containing at least 1
 CC restriction enzyme site, a transcription terminator (pref. the dimer
 CC comprising the strands given in AAQ89697-98) and a drug resistance
 CC marker.
 XX
 SQ Sequence 94 BP; 20 A; 28 C; 26 G; 20 T; 0 other;
 XX

Query Match 1.5%; Score 36; DB 16; Length 94;
 Best Local Similarity 97.9%; Pred. No. 1.2e+02;
 Matches 47; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 ttggtaccagctcgatcc-ctagtaacggccgcagctgctggaa 47
 Db 4 ttggtaccagctcgatccactagtagtaacggccgcagctgctggaa 51
 |||||
 |||||

RESULT 13
 AAV37197/c
 ID AAV37197 standard; DNA; 80 BP.
 XX
 AC AAV37197;
 XX
 DT 14-SEP-1998 (first entry)
 XX
 DE Oligonucleotide sequence of the specification.
 XX
 KW Genotype; phenotype; molecular evolutionary engineering;
 KW functional biopolymer; virus; ss.
 XX
 OS Synthetic.
 XX
 PN WO9816636-A1.
 XX
 PD 23-APR-1998.
 XX
 PF 17-OCT-1997; 97WO-JP03766.
 XX
 PR 17-OCT-1996; 96JP-0274855.
 XX
 PA (MITU) MITSUBISHI CHEM CORP.
 XX
 PI Fusimi Y, Miyamoto E, Nemoto N, Yanagawa H;
 XX
 DR WPI; 1998-261039/23.
 XX

PT Virus containing nucleic acid and protein sections - for use in
 PT modification and creation of functional bio:polymers such as
 PT enzymes, antibodies and ribozyme(s)
 XX
 PS Example 1; Page 44; 68pp; Japanese.
 XX
 CC The present sequence is used in the course of the invention. The
 CC specification describes a molecule for bringing together genotype with
 CC phenotype (in vitro virus). The molecule contains a nucleic acid
 CC fragment having a base sequence corresponding to a genotype, covalently
 CC bonded to a protein fragment containing a protein participating in
 CC phenotype expression, the 3'-end of the nucleic acid part being bonded
 CC to the C-terminus of the protein part via a puromycin moiety. The nucleic
 CC acid fragment preferably consists of RNA corresponding to the gene (free
 CC from a termination codon), a spacer (such as two-stranded DNA), a peptide
 CC adapter (such as a DNA-trNA hybrid containing an anticodon corresponding
 CC to the terminator codon of the gene) and a puromycin-containing cap
 CC capable of binding to an amino acid residue. Translation of the virus is
 CC performed in a non-cellular (preferably ribosomal) system (e.g. E. coli
 CC ribosome), the protein synthesised by the translation attaching to the
 CC puromycin cap, resulting in the complete in vitro virus structure. The
 CC method is used in molecular evolutionary engineering to optimise function
 CC of a functional biopolymer such as an enzyme, antibody or ribozyme, or
 CC to generate new functionality. The virus can be used for the functional
 CC optimisation of nucleic acid or protein sequences.
 XX
 SQ Sequence 80 BP; 7 A; 10 C; 1 G; 62 T; 0 other;
 XX

Query Match 1.5%; Score 35.8; DB 19; Length 80;
 Best Local Similarity 78.2%; Pred. No. 1.2e+02;
 Matches 43; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 QY 2311 totttaataaaatagagtgattttaaacttaaaaaaAAAAAAAAAAAAAAAAAAAA 2365
 Db 68 TCATTAATAA 14
 |||||
 |||||

RESULT 14
 AAV05720
 ID AAV05720 standard; cDNA; 85 BP.
 XX
 AC AAV05720;
 XX
 DT 05-JUN-1998 (first entry)
 XX
 DE Nucleotide sequence of the 3' portion from clone Az302_1.
 XX
 KW Colon; secreted protein; EST sequence; homology; antibody;
 KW immunoassay reagent; nutritional supplement; therapeutic activity; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO9746683-A2.
 XX
 PD 11-DEC-1997.
 XX
 PF 06-JUN-1997; 97WO-US09878.
 XX
 PR 07-JUN-1996; 96US-0659224.
 XX
 PA (GENY) GENETICS INST INC.
 XX
 PI Bowman M, Evans C, Jacobs K, Lavallie ER, McCoy JM;
 PI Merberg D, Racie LA, Spaulding V, Treacy M;
 XX
 DR WPI; 1998-042191/04.
 XX
 PT Nucleic acids encoding secreted proteins from clones within ATCC
 PT 98076 - useful as immuno-modulators, anti-proliferative agents,
 PT regulators of cell differentiation and tissue growth, etc
 XX
 SQ Claim 13; Page 57; 99pp; English.
 XX

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 24, 2002, 19:08:58 ; Search time 2466.96 Seconds
(without alignments)
12939.123 Million cell updates/sec

Title: US-09-923-831-42

Perfect score: 2365

Sequence: 1 ttgttaccgagctggatcc.....aaaaaaaaaaaaaaaaaaaaa 2365

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 297742

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estau:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_gss:*
- 13: em_gss_hum:*
- 14: em_gss_inv:*
- 15: em_gss_pln:*
- 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	48.2	2.0	79	9	AA781190 aj24f04.s
C 2	44.4	1.9	86	9	AI701097 wc78a09.x
C 3	43	1.8	91	10	C25755 C25755 Dict
4	42.6	1.8	96	10	R28767 F0-233D 22
C 5	42	1.8	83	9	AI499570 to02g10.x
6	42	1.8	89	9	AU053810 AU053810
7	42	1.8	93	9	AI945829 bs17d12.y
8	41.8	1.8	90	9	AU054081 AU054081
9	41.6	1.8	96	10	B1941603 sc77a05.y
C 10	41.4	1.8	75	9	AI462919 vb82g08.x
11	41.4	1.8	94	9	AA241234 AA241234
12	41.4	1.8	99	9	AU051957 AU051957
C 13	41	1.7	86	9	AI636507 ts91b09.x
C 14	41	1.7	93	10	BG122005 BG122005
C 15	41	1.7	98	10	BE319731 BE319731
16	40.6	1.7	96	9	AA120449 AA120449
C 17	40.4	1.7	92	9	AA554497 ni36d04.s

C 18	40.4	1.7	99	9	AI638644
19	40.4	1.7	100	9	AI210082
C 20	40.2	1.7	73	9	AI050700
C 21	40.2	1.7	87	9	AW085350
22	40.2	1.7	95	10	BF338778
C 23	40.2	1.7	100	9	AA607932
C 24	39.8	1.7	73	10	R43518
C 25	39.8	1.7	97	9	AI628325
C 26	39.6	1.7	100	9	AA624410
C 27	39.4	1.7	64	9	AI955977
C 28	39.4	1.7	68	10	N21402
29	39.4	1.7	80	10	BI142670
C 30	39.4	1.7	83	9	AI061323
C 31	39.4	1.7	95	9	AI678446
C 32	39.4	1.7	97	9	AU038741
C 33	39.4	1.7	100	9	AI355779
34	39.2	1.7	75	9	AI938927
35	39.2	1.7	80	12	CNS01TV2
36	39.2	1.7	97	10	C84172
37	39	1.6	83	9	AU052639
C 38	39	1.6	87	9	AI419826
C 39	39	1.6	88	10	BF635188
C 40	39	1.6	90	9	AI439664
41	39	1.6	93	9	AU038793
C 42	39	1.6	98	9	AW866658
C 43	39	1.6	99	9	AI611720
C 44	38.8	1.6	72	9	AW148882
C 45	38.8	1.6	78	2	HSW010524

ALIGNMENTS

RESULT 1

AA781190/c

LOCUS

DEFINITION

aj24f04.s1 Soares-testis_NHT Homo sapiens CDNA clone 1391263 3'

Similar to TR:P96614 P96614 ATP-DEPENDENT RNA HELICASE DEAD

HOMOLOG. ; mRNA sequence.

ACCESSION

AA781190

VERSION

AA781190.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 79)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaudo

, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Possible reversed clone: similarity on wrong strand

Insert Length: 708 Std Error: 0.00

Seq primer: -40m13 fwd. Et from Amersham

High quality sequence stop: 1.

Location/Qualifiers

1..79

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="1391263"

/sex="male"

FEATURES

source

```

/lab_host="DH109"
/notes="Vector: p77T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTACCATCTGAGTGGAGGCGCCAAATTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p77T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
25 a 17 c 13 g 24 t

```

```

BASE COUNT
ORIGIN

```

```

Query Match      2.0%; Score 48.2; DB 9; Length 79;
Best Local Similarity 76.6%; Pred No. 1.le+03;
Matches 59; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1853 gagactgagtcacatgacgttacacatgtctataattgttgcattccaggaattg 1912
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 78 GAGGCTTGATGATTGATGTCACACATGATATATATGATGATCTCCAGGGAATATG 19
QY 1913 acaatacgtacacga 1929
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 18 ACATATGTCACAGA 2

```

```

RESULT 2
LOCUS      AI701097/c
DEFINITION wc78h09.x1 NCI-CGAP.Pan1 Homo sapiens cDNA clone IMAGE:2324801 3'
            similar to contains TAR1.b2 MSRL repetitive element ;, mRNA
            sequence.
ACCESSION  AI701097
VERSION     AI701097.1 GI:4988937
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 86)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-re@mail.nih.gov
            Life Technologies catalog #: 11548-013
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/BLNL at:
            www-bio.llnl.gov/obrp/image/image.html
            Insert Length: 1789 Std Error: 0.00
            Seq primer: -400P from Gibco
            High quality sequence stop: 69.
FEATURES   Location/Qualifiers
            1..86
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:2324801"
            /clone_lib="NCI-CGAP.Pan1"
            /tissue_type="adenocarcinoma"
            /lab_host="DH10B"
            /note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI;
            Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.72 kb. Life Technologies catalog #:
            11548-013"

```

```

BASE COUNT      17 a 3 c 2 g 64 t
ORIGIN

```

```

Query Match      1.9%; Score 44.4; DB 9; Length 86;

```

```

Best Local Similarity 73.1%; Pred No. 3.8e+03;
Matches 57; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 2288 atgaatccagtggtttatctattctttaaataaaatagaagtatttaaaccttaaaaaa 2347
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 AAAAAACCGGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 26
QY 2348 aaaaaaataaaaaa 2365
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25 AAAAAAATAAAAAA 8

```

```

RESULT 3
LOCUS      C25755
DEFINITION C25755 Dictyostelium discoideum FC (H.Urushihara) Dictyostelium
            discoideum cDNA Clone FC-A208, mRNA sequence.
ACCESSION  C25755
VERSION     C25755.1 GI:2292541
KEYWORDS   EST.
SOURCE     Dictyostelium discoideum.
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
ORGANISM   Dictyostelium discoideum.
            Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

```

```

REFERENCE  1 (bases 1 to 91)
            Suzuki,K., Shimizu,H. and Urushihara,H.
            Sexual cDNA in D. discoideum(970724)
            Unpublished (1997)
            Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
            Email: d402huesakura.cc.tsukuba.ac.jp.
FEATURES   Location/Qualifiers
            1..91
            /organism="Dictyostelium discoideum"
            /strain="KAX3"
            /db_xref="taxon:44689"
            /clone="FC-A208"
            /clone_lib="Dictyostelium discoideum FC (H.Urushihara)"

```

```

BASE COUNT      59 a 0 c 5 g 27 t
ORIGIN

```

```

Query Match      1.8%; Score 43; DB 10; Length 91;
Best Local Similarity 73.3%; Pred. No. 6e+03;
Matches 55; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
QY 2291 gaatccagtggtttatctattctttaaataaaatagaagtatttaaaccttaaaaaa 2350
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 10 GATGTGTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 69
QY 2351 aaaaaaataaaaaa 2365
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 70 AAAAAAATAAAAAA 84

```

```

RESULT 4
LOCUS      R28767
DEFINITION R28767 F0-233D 22 week old human fetal liver cDNA library Homo sapiens
            cDNA clone F0-233D 5', mRNA sequence.
ACCESSION  R28767
VERSION     R28767.1 GI:6514139
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE  1 (bases 1 to 96)
            Choi,S.S., Yun,J.W., Choi,E.K., Cho,Y.G., Sung,Y.C. and Shin,H.S.
            Construction of a gene expression profile of a human fetal liver by
            single-pass cDNA sequencing
            Mamm. Genome 6 (9), 653-657 (1995)
JOURNAL

```


4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estevatson.wustl.edu
 This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Hunttsville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: ccu@resgen.com
 Seq primer: -40RP from Gibco
 High quality sequence stop: 87.

FEATURES

source

```

1..96
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl018-898"
/clone_lib="Gm-cl018"
/tissue_type="leaves of greenhouse grown plants"
/dev_stage="2-3 weeks old"
/lab_host="DH10B (Gibco BRL)"
/notes="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from 2-3 week old greenhouse grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker and Dr. John Expelding."

```

BASE COUNT 66 a 6 c 10 g 14 t

ORIGIN

```

Query Match 1.8%; Score 41.6; DB 10; Length 96;
Best Local Similarity 67.0%; Pred. No. 9.5e+03;
Matches 59; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 2277 tgtttgaatatagatcagtcagtttttatacttttttaataaaatagaatttttaa 2336
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 9 TATCTAAATGAGAAATATTCTAGTACTTGTGGAAAAAATAAAAAAATAAAAAA 68

QY 2337 acttaaaaaaataaaaaaataaaaaa 2364
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 69 AAAAAAATAAAAAAATAAAAAAATAAAAAA 96

```

```

RESULT 10
AI462919/c
LOCUS
DEFINITION AI462919 75 bp mRNA linear EST 09-MAR-1999
IMAGE:763550 3', mRNA sequence.
ACCESSION AI462919
VERSION AI462919.1 GI:4316949
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 75)
Marras, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person
, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schuck, R., Ritter
, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999
Unpublished (1999)
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:464470
 This clone was previously sequenced on the 5' end only, this new
 data is from the 3' end
 High quality sequence stop: 63.

FEATURES

source

```

1..75
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:763550"
/clone_lib="Soares mouse 3NME12 5"
/sex="unknown"
/tissue_type="fetus"
/dev_stage="12.5dpc total fetus"
/lab_host="DH10B"
/notes="Organ: whole fetus; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAGTGGAGGCGCCCTATTTTATTTTATTTT 3'], on total mouse RNA [provided by Minoru Ko, Wayne State Univ.]; double-stranded cDNA was ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

```

BASE COUNT 20 a 0 c 0 g 55 t

ORIGIN

```

Query Match 1.8%; Score 41.4; DB 9; Length 75;
Best Local Similarity 76.1%; Pred. No. 1.1e+04;
Matches 51; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2299 tgtttatacttttttaataaaatagaatttttaaaacttaaaaaaataaaaaa 2358
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 72 TTTTITTTTTTTTTTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 13

QY 2359 aaaaaa 2365
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 12 AAAAAAA 6

```

```

RESULT 11
AA241234
LOCUS
DEFINITION AA241234 94 bp mRNA linear EST 04-MAR-1997
IMAGE:653338 5', mRNA sequence.
ACCESSION AA241234
VERSION AA241234.1 GI:1864568
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 94)
Marras, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

```

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:399186

Putative full length read

vector to vector length is 117

Seq primer: -28ml3 rev2 ET from Amersham.

FEATURES

Location/Qualifiers

1..94

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:653338"

/clone_lib="Soares mouse lymph node NbMLN"

/sex="male"

/tissue_type="lymph node"

/dev_stage="4 weeks"

/lab_host="DH10B"

/note="Organ: lymph node; Vector: p7T3D-Pac (Pharmacia)

with a modified polylinker; Site_1: Not I; Site_2: Eco RI;

1st strand cDNA was primed with a Not I - oligo(dT) primer

15.

TGTTACCAACTGAAGTGGAGCGCGGATCTTTTCTTTTCTTTTCTTTTCTTTT

3']; double-stranded cDNA was ligated to Eco RI adaptors

(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified p7T3 vector. RNA

provided by Dr. Bertrand Jordan. Library constructed and

normalized by Bento Soares and M. Fatima Bonaldo."

46 a 10 c 15 g 23 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 1.8%; Score 41.4; DB 9; Length 94;

Matches 51; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2298

ggttttactcttttataaataagtagatttataaacttaaaacttaaaaaa 2357

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 28 GTGTTGCTCTCTTTGTTTAATAAATATGACGTCAGTCAAAAAA 87

QY 2358

aaaaaa 2364

|||||

Db 88

AAAAAA 94

RESULT 12

AU051957

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

1..99

/organism="Dictyostelium discoideum"

/strain="AX4"

/db_xref="taxon:44689"

/clone_lib="Dictyostelium discoideum SS (H. Urushihara)"

FEATURES

Location/Qualifiers

1..99

/organism="Dictyostelium discoideum"

/strain="AX4"

/db_xref="taxon:44689"

/clone_lib="Dictyostelium discoideum SS (H. Urushihara)"

BASE COUNT
ORIGIN

73 a 1 c 2 g 23 t

Query Match

Best Local Similarity 1.8%; Score 41.4; DB 9; Length 99;

Matches 51; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 2299

tgttttactcttttataaataagtagatttataaacttaaaacttaaaaaa 2358

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 16 TTTTCTTTTCTTTTATAAAAAA 75

QY 2359

aaaaaa 2365

|||||

Db 76

AAAAAA 82

RESULT 13

AI636507/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..86

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2238617"

/clone_lib="NCI CGAP GC6"

/tissue_type="pooled germ cell tumors"

/lab_host="DH10B"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified

polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA

from the normalized library NCI CGAP GC4 was prepared, and

ss circles were made in vitro. Following HAP purification,

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from a pool

of 5,000 clones made from the same library (clonoids

1257096-1258631, 1469064-1470983, and 1475592-1476743).

Subtraction by Bento Soares and M. Fatima Bonaldo."

12 a 6 c 9 g 59 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 1.7%; Score 41; DB 9; Length 86;

Matches 50; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 2301

ttttactcttttataaataagtagatttataaacttaaaaaa 2360

```

||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 69 TTTTGGCTTTTATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 10

QY 2361 aaaa 2365
|||||
Db 9 AAAAA 5

RESULT 14
LOCUS BG122005/c
DEFINITION BG122005.1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4449713 5',
mRNA sequence.
ACCESSION BG122005
VERSION BG122005.1 GI:12615514
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 93)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L9600234 row: b column: 18
High quality sequence stop: 93.
FEATURES
source
1..93
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4449713"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 20 a 8 c 10 g 55 t
ORIGIN
Query Match 1.7%; Score 41; DB 10; Length 93;
Best Local Similarity 76.9%; Pred. No. 1.2e+04;
Matches 50; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 2301 ttctacactttcttaaaatagaagtatttaaaccttaaaaaaataaaaaa 2360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 90 TTTTATTTTATTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 31

QY 2361 aaaa 2365
|||||
Db 30 AAAAA 26

RESULT 15
LOCUS BE319731/c
DEFINITION BE319731 NF020E07RT1F1051 Developing root Medicago truncatula cDNA clone
ACCESSION BE319731
VERSION NF020E07RT 5', mRNA sequence.
KEYWORDS EST.

```

```

SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 98)
REFERENCE Watson, B.S., Shin, H.-S., Lopez-Meyer, M., Scott, A.D., Harris, A.R.,
AUTHORS Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W., May,
G.D. and Paiva, N.L.
Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula root library
JOURNAL Unpublished (2000)
COMMENT On Jul 14, 2000 this sequence version replaced gi:9193508.
Contact: Paiva NL
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7317
Fax: 580 221 7380
Email: nlpaiva@noble.org
Medicago Genome Initiative accession: MGI:S:16962
Insert Length: 721 Std Error: 0.00
Plate: 020 row: E column: 07
Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES
source
1..98
Location/Qualifiers
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF020E07RT"
/clone_lib="Developing root"
/tissue_type="root"
/dev_stage="Pooled developmental"
/note="Vector: Lambda zap; Total RNA was extracted from
non-modulated roots of plants grown in 1 mM nitrate
medium. Samples were taken at four time points
(approximately two days, one, two and six weeks post
germination) representing early seedling growth to
nitrogen limitation."
BASE COUNT 15 a 6 c 0 g 77 t
ORIGIN
Query Match 1.7%; Score 41; DB 10; Length 98;
Best Local Similarity 76.9%; Pred. No. 1.2e+04;
Matches 50; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 2301 ttctacactttcttaaaatagaagtatttaaaccttaaaaaaataaaaaa 2360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97 TTTTATTTTATTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 38

QY 2361 aaaa 2365
|||||
Db 37 AAAAA 33

Search completed: September 24, 2002, 19:56:52
Job time: 2874 sec

```
